



Infectología en la era de la genómica

Andrea I. Moreno Switt y Viviana Toledo

Infectious diseases in the genomic era

Next generation sequencing (NGS) technologies have arrived, changing research and infectious disease research into a new era, the “genomic era”. Currently, the developed world is introducing NGS in a number of applications, including clinical diagnostics, epidemiology, and microbiology. In developing countries NGS is being progressively introduced. Technologies currently available allow to sequence the whole genome of bacterial and viral strains for an approximate cost of \$100 USD, which is highly cost savings compared to old-technologies for genome sequencing. Here we review recent publication of whole genome sequencing used for, (i) tracking of foodborne outbreaks, with emphasis in *Salmonella* and *Listeria monocytogenes*, (ii) building genomic databases for Governments, (iii) investigating nosocomial infections, and (iv) clinical diagnosis. The genomic era is here to stay and researchers should use these “massive databases” generated by this technology to decrease infectious diseases and thus improve health of humans and animals.

Key words: Genomics, molecular epidemiology, next generation sequencing.

Palabras clave: Genómica, epidemiología molecular, nueva generación de secuenciadores.

Facultad de Ecología y Recursos Naturales.
Escuela de Medicina Veterinaria
Universidad Andrés Bello
Santiago, Chile.

Los autores declaran no tener conflictos de interés.
Este manuscrito fue financiado por Proyecto Fondecyt de iniciación 11140108 a nombre de Andrea Moreno Switt. Viviana Toledo es financiada por la Vicerrectoría de Investigación y Doctorado de la Universidad Andrés Bello.

Recibido: 27 de mayo de 2015
Aceptado: 18 de agosto de 2015

Correspondencia a:

Andrea I Moreno Switt
Andrea.moreno@unab.cl

Evolución de la secuenciación masiva: desde los primeros secuenciadores al secuenciador de bolsillo

El ADN está compuesto por cuatro letras que son la base de la vida y forman parte del código genético. A través del tiempo, son muchas las tecnologías que se han desarrollado para descifrar este código; en los años 70, Sanger introdujo por primera vez un método, que permitió identificar de forma manual secuencias nucleotídicas basándose en el proceso de la replicación¹. A pesar que el método de Sanger ha mejorado desde sus inicios al desarrollar un sistema de secuenciación automático¹, el costo por secuenciación con esta tecnología sigue siendo extremadamente caro y demorado; por ejemplo, el costo de secuenciar el genoma humano en el año 2003 fue de 3 billones de dólares, el cual logró completarse en 13 años^{1,2}. Posteriormente, un hito en la secuenciación fue el desarrollo de la secuenciación de nueva generación (“next generation sequencing”-NGS), definición utilizada para describir tecnologías que permiten secuenciar de manera masiva a un precio mucho más económico que el método Sanger y a una velocidad mucho mayor^{1,3}. Dentro de las tecnologías que pertenecen a la NGS se encuentran los equipos MySeqTM de Illumina, el 454TM de Roche y el IonTorrentTM de Life Technologies⁴. A pesar que la química de secuenciación es diferente para cada equipo (ej. piro-secuenciación o secuenciación por síntesis); todas han logrado disminuir exponencialmente el costo de secuenciar^{3,5}. La evolución en la secuenciación no sólo se manifiesta por el desarrollo de estas nuevas

tecnologías, sino además en la preparación del ADN que se va a secuenciar. Para secuenciar un genoma con Sanger, es necesario realizar una genoteca (librería genética) *in vivo*, que implica clonar los segmentos de ADN en un plásmido, para luego introducirlos a una cepa bacteriana, posteriormente amplificar los fragmentos por reacción de polimerasa en cadena (PCR) y finalmente purificar el amplicón. Todo este proceso es extremadamente laborioso y además muy costoso¹. En cambio, en la NGS la genoteca se realiza *in vitro* donde el ADN genómico se fragmenta en trozos a los que posteriormente se les adicionan los adaptadores (secuencias conocidas) que permitirán realizar inmediatamente una amplificación por PCR de los fragmentos, utilizando partidores que son específicos para los adaptadores ligados (ver figura en Shendure y Ji, 2008 por más detalles)⁵.

Cuando se visualiza la evolución de la NGS se vislumbra un cambio de paradigma en cómo se diagnostican las enfermedades infecciosas y a su vez cómo está cambiando la epidemiología. Equipos como el HiSeq de Illumina permiten hoy en día secuenciar un genoma bacteriano por aproximadamente \$100 USD en aproximadamente 10 h⁴. Aún más, tal es la evolución de la NGS que desde el año 2012 la empresa británica Oxford Nanopore (<https://www.nanoporetech.com>) anuncia la creación del MinIONTM, un secuenciador del tamaño de un *pendrive*, que a pesar de no estar aún en el mercado, es una tecnología portátil que permite secuenciar en tiempo real una bacteria con una cobertura de 30 X (30 veces) en una hora^{6,7}. Todos estos avances en la NGS determinan que hoy en día estamos viviendo la era de la genómica, evidenciada por un au-



mento exponencial en los genomas bacterianos y virales secuenciados, por un mejoramiento en los algoritmos de análisis y por un incremento en la aplicabilidad de la secuenciación en el estudio de enfermedades infecciosas⁴.

Secuenciación como herramienta para rastrear brotes, la transformación de la epidemiología

Los métodos moleculares de sub-tipificación que permiten discriminar entre cepas bacterianas involucradas presuntivamente en un brote, son herramientas que han ayudado a mejorar de manera significativa la capacidad de rastrear brotes^{4,8,9}. Actualmente, el método “estándar de oro” de subtipificación bacteriana es la electroforesis de campo pulsado (PFGE por su sigla en inglés)⁴. Sin embargo, el poder de discriminación del PFGE, es decir, su capacidad de diferenciar entre cepas, es limitada si se compara con la secuenciación de genomas completos. De hecho, se estima que sólo 1% del genoma es analizado cuando se observa el patrón de bandas generadas por el PFGE^{10,11}. En varios estudios e investigaciones epidemiológicas ha quedado de manifiesto la limitación del PFGE, sobre todo en casos donde las cepas involucradas son de tipo monomórficas o clonales. Por ejemplo, para algunos serotipos de *Salmonella* la sub-tipificación con PFGE no es útil, ya que no discrimina entre cepas muy clonales¹²; es el caso de *Salmonella* Enteritidis, el serotipo más común de este patógeno. Cuando las cepas de *S. Enteritidis* se discriminan con PFGE, entre 40 y 50 % de las cepas presentan el mismo patrón^{13,14}; por lo tanto, realizar la investigación epidemiológica, rastrear la fuente e identificar la ruta de transmisión se convierte en un proceso altamente complejo y la utilidad de PFGE, es más bien dudosa. Ejemplo de otros patógenos bacterianos en que la PFGE no es de mucha utilidad incluyen *Staphylococcus aureus*, *Clostridium difficile*, *Klebsiella pneumoniae* y *Mycobacterium tuberculosis*¹⁴. Recientemente, den Bakker y cols.¹⁴ reportaron un análisis filogenético derivado de la identificación de polimorfismos en los nucleótidos en el genoma completo de 93 cepas de *S. Enteritidis*, metodología que ellos denominaron “*whole-genome cluster analysis*”. Los autores demostraron la superioridad de este análisis en comparación con la PFGE, ya que por medio de un estudio retrospectivo y uno prospectivo fueron capaces de identificar las cepas involucradas en un brote y diferenciarlas de las cepas que interferían con las investigaciones epidemiológicas. Por otro lado, también se describe que, gracias a la secuenciación, se han descubierto nuevas fuentes de contagio existentes para *Listeria monocytogenes*. En E.U.A., la Food and Drug Administration (FDA) encontró alimentos asociados a brotes de listeriosis en fuentes que no se asociaban comúnmente a la bacteria, como son

las manzanas acarameladas (<http://www.cdc.gov/amdstories/listeria-caramel-apples.html>).

En la actualidad, en algunos países desarrollados ya se están secuenciando de manera frecuente genomas de bacterias y virus que son aislados de pacientes, del ambiente, alimentos y animales. A la fecha de agosto de 2014, el Departamento de Salud Pública de Inglaterra había secuenciado 28.000 bacterias y virus (ej. VIH, *M. tuberculosis*, virus de la hepatitis C), como parte de un proyecto cuyo objetivo es secuenciar un total de 100.000 patógenos¹⁵.

De igual manera, en E.U.A., desde el año 2013, todas las cepas de *L. monocytogenes* provenientes de casos clínicos de listeriosis son secuenciadas en los Centers for Disease Control and Prevention (CDC). (<http://www.cdc.gov/media/releases/2013/p0604-listeria-poisoning.html>); en esta misma línea, este organismo además se propuso como desafío comenzar a utilizar métodos independientes al cultivo para realizar las investigaciones epidemiológicas. Una de ellos es la Metagenómica (ver abajo para más detalles) (<http://www.cdc.gov/pulsenet/next-generation.html>).

Otras instituciones de E.U.A., también están desarrollando proyectos para la investigación de enfermedades zoonóticas transmitidas por los alimentos. Existen actualmente dos grandes proyectos; uno de la Universidad de California, proyecto titulado “100K Foodborne Pathogen Genome Project” (<http://100kgenome.vetmed.ucdavis.edu/index.cfm>), el que secuenciará el genoma de 100.000 patógenos que son agentes causales frecuentes de enfermedades transmitidas por los alimentos (ETA) (ej. *Salmonella*, *Listeria*). EL objetivo de este proyecto es almacenar información para facilitar el diagnóstico de las ETAs y desarrollar soluciones que permitan mejorar la salud pública (<http://100kgenome.vetmed.ucdavis.edu/index.cfm>).

El otro gran proyecto existente, es el “GenomeTrakr” del FDA, el que ha hasta la fecha (Mayo, 2015) ha secuenciado 12.000 cepas de *Salmonella* y 3.100 cepas de *L. monocytogenes*. Este proyecto secuencia mensualmente alrededor de 800 bacterias, lo que permite generar información que es utilizada para rastrear brotes de manera más rápida, además de crear una base de datos de cepas aisladas de distintas fuentes y orígenes que en el futuro van a permitir rastrear el origen de un brote más fácilmente. Para mayor información ver: <http://www.fda.gov/Food/FoodScienceResearch/WholeGenomeSequencingProgramWGS/ucm403550.htm>.

En teoría, la generación de una base de datos con miles de genomas obtenidos alrededor del planeta sobre todo en el mundo globalizado en que estamos, en que consumimos condimentos importados desde Asia, carnes de Brasil, chocolate de Europa, etc., podrá en el futuro facilitar la identificación de una fuente de contaminación. En ese



sentido, este proyecto de la FDA publicó recientemente el anuncio que 33 cepas de *Salmonella* aisladas en Chile fueron secuenciadas para entender la diversidad filogenética de cepas de *Salmonella* en Sudamérica¹⁶. Demostrando esto, que ésta ya es una realidad nacional.

En brotes intrahospitalarios donde la mejor forma de prevención es evitar la transmisión persona-persona, la utilización de secuenciación ha demostrado ser muy efectiva^{8,17}. Por ejemplo, en un brote de *Staphylococcus aureus* resistente a meticilina (SARM) ocurrido en la unidad de cuidados intensivos neonatal del Cambridge University Hospitals se secuenciaron los genomas de siete cepas de SARM asociadas al brote y siete cepas asociadas a portadores o a casos de bacteriemia, todos pertenecientes al mismo hospital. Para los análisis, los autores utilizaron una metodología similar a la descrita por den Bakker y cols., (ver más arriba) donde las substituciones de nucleótidos en los genes del genoma-*core* (genes presentes en todas las cepas de *S. aureus*) fueron utilizadas para realizar el análisis filogenético y así identificar las cepas del brote y los eventos de transmisión intrahospitalaria (infantes-madres-personal del hospital) que no habían podido ser identificados. Además, los autores predijeron *in silico* (detección a través de algoritmos de computación) el perfil de resistencia antimicrobiana, el cual fue después validado por el antibiograma⁸.

Las NGS no sólo se han utilizado en casos en que el agente causal son bacterias; también se ha utilizado para el estudio de brotes causados por virus, donde cobran especial relevancia ya que éstos son difíciles de cultivar y los métodos actualmente utilizados como la serología o la secuenciación de un gen marcador son bastante limitados en su poder de discriminación entre cepas; además estas técnicas no permiten elucidar las dinámicas evolutivas de estos virus, ni menos sus patrones de transmisión^{17,18}. Por esto, cuando en el año 2009 ocurrió la pandemia del virus influenza H1N1, Baillie y cols.¹⁹, secuenciaron 153 cepas de los virus aislados de pacientes en el Reino Unido para elucidar su dinámica de evolución. Los autores encontraron que existían múltiples linajes del virus, incluyendo 13 linajes que eran exclusivos del Reino Unido; además describieron la persistencia de algunos linajes en la población. Un ejemplo más reciente es la secuenciación de cepas del virus del Ébola asociadas al brote más grande nunca antes visto; en febrero de 2014 se confirmaron los primeros casos en África y para elucidar el origen de este brote, un grupo multidisciplinario de investigadores de distintas nacionalidades, secuenciaron los genomas de 99 virus aislados de 78 pacientes distintos²⁰. En este estudio, los autores identificaron un gran número de mutaciones en los virus que ellos secuenciaron, mutaciones que podrían afectar, por ejemplo, los sistemas moleculares de detección (ej., partidores no amplificarían en ADN de estos virus mutados), los virus secuenciados de pacientes

en Sierra Leona presentaban genomas que no serían detectados por muchos de los ensayos moleculares²⁰. Por último, en este estudio la secuenciación permitió además predecir que un solo evento zoonótico de transmisión podría haber sido el origen de este brote.

En un brote, no solamente basta con identificar la cepa involucrada, sino además es necesario diseñar un método veloz que facilite la posterior identificación de la cepa en nuevos casos clínicos o en una fuente de transmisión. Fue así como en un brote de *Escherichia coli* O104:H4 en Europa se secuenciaron cepas asociadas al brote simultáneamente por distintos grupos de investigación²¹. Esta cepa presentaba características únicas, ya que era productora de la toxina Shiga, pero pertenecía a un serotipo que nunca antes había sido asociado a cuadros del síndrome hemolítico urémico (SHU)²¹. La secuenciación de las cepas no sólo permitió explicar el fenotipo presente (inserción del bacteriófago con los genes de la toxina Shiga), sino además facilitó el desarrollo de un ensayo diagnóstico de qPCR) específico para la cepa involucrada en el brote²¹⁻²³.

La epidemiología de hoy en día se ve desafiada por el aumento exponencial de información, cientos a miles de genomas que se han secuenciado, junto con los que estarán disponibles públicamente en bases de datos como GenBank en el futuro cercano. Es así que para facilitar el procesamiento de cantidades monumentales de información, en Dinamarca se está desarrollando un proyecto que tiene como objetivo crear soluciones en la red para el análisis de las secuencias. El proyecto se denomina “Center for Genomic Epidemiology” (<http://www.genomicepidemiology.org/>), y en este sitio se pueden analizar, por ejemplo el “resistoma” (conjunto total de genes de resistencia), o el “viruloma” (conjunto total de genes de virulencia), como también realizar análisis filogenéticos en genomas ensamblados (lecturas obtenidas del secuenciador ya ensambladas), entre otras aplicaciones. Muchas de las instituciones internacionalmente conocidas (USDA, FDA, WHO, otras) están de acuerdo en que en unos cinco a 10 años más la secuenciación va a ser la herramienta estándar para identificar brotes y realizar estudios epidemiológicos. Con equipos como el MinION™ podremos en un futuro cercano por ejemplo, secuenciar una muestra de sangre de un paciente con septicemia, identificar el agente causal y predecir el perfil de resistencia en tan solo unas horas.

Aplicaciones de metagenómica en el diagnóstico de enfermedades zoonóticas y el descubrimiento de nuevos virus

La metagenómica se refiere al estudio del material genético presente en una muestra que puede ser agua, sangre, fecas, etc.²⁴. En este tipo de estudios se extrae el ADN total de una muestra, se secuencian los genes extraídos



y, por medio de la bio-informática, se identifican las diferentes unidades taxonómicas presentes en la muestra⁴.

La metagenómica es independiente del cultivo, lo que la hace una técnica muy atractiva para identificar patógenos virales o bacterianos que no crecen fácilmente en el laboratorio²⁵. Recientemente se utilizó la metagenómica en un hospital en Alemania para identificar el agente causal de neumonía grave en pacientes que dieron negativos a todos los patógenos que comúnmente se asocian a neumonía²⁶. En este estudio, el ADN fue purificado de muestras de lavado broncoalveolar, el que fue posteriormente secuenciado con el MiSeq™ de Illumina; en un total de 50 h (incluyendo los análisis bio-informáticos), basándose en los datos obtenidos, los investigadores fueron capaces de identificar a *Chlamydophila psittacii* como agente causal de las neumonías por lo cual los médicos a cargo pudieron modificar el tratamiento, lo que se tradujo en la recuperación de uno de los pacientes²⁶. De la misma manera, la metagenómica se ha utilizado para diagnosticar la presencia de *Campylobacter jejuni*²⁷, norovirus¹⁷, inclusive la presencia de parásitos muy extraños (*Kudoa septempunctata*)²⁸, en pacientes con sintomatología de diarrea.

Históricamente, las enfermedades infecciosas han sido diagnosticadas utilizando técnicas que fueron desarrolladas a fines del 1800, como son la tinción de Gram y el aislamiento del agente causal en medios de cultivo²⁹. Últimamente, el avance en las técnicas moleculares ha hecho que además se esté incorporando la PCR o la qPCR para el diagnóstico de enfermedades infecciosas; sin embargo su uso hasta el día de hoy es más común en investigación que en diagnóstico clínico²⁹. Una de las razones por las cuales se ha incorporado la metagenómica al diagnóstico de agentes infecciosos es el hecho que las técnicas de microbiología clásica o detección por PCR no pueden identificar la presencia de patógenos desconocidos²⁹. Muchos de los patógenos desconocidos son de origen animal; por ejemplo, el CDC de E.U.A. estima que, aproximadamente 60-80% de los casos de enfermedades zoonóticas transmitidas por los alimentos son causadas por patógenos desconocidos³⁰. A eso se suma que 75% de las enfermedades emergentes son de origen zoonótico y que 60% de las enfermedades infecciosas en humanos son además zoonóticas (<http://www.cdc.gov/24-7/cdcfastfacts/zoonotic.html>). Es así que se han descubierto muchos virus en animales silvestres utilizando metagenómica; además se han desarrollado algoritmos que pueden predecir el riesgo de transmisión a humanos o de incluso calcular el riesgo de una pandemia, todo esto con el fin de poder utilizar la información generada de la metagenómica de una forma racional^{25,29}.

En el futuro, el uso de la metagenómica en el diagnóstico de enfermedades infecciosas podría ser una prueba de rutina en muchos laboratorios de microbiología clínica.

En teoría, la información generada de la secuenciación podría utilizarse para identificar no sólo el agente causal, sino predecir el patrón de resistencia o incluso podría utilizarse para estudiar el efecto del tratamiento. Sin embargo, a pesar de todas las herramientas actuales, aún estaría en discusión el tema ético del uso de metagenómica en el diagnóstico de enfermedades infecciosas, principalmente por el hecho que al secuenciar “todo” el ADN de la muestra, el ADN del paciente podría ser secuenciado y esa información podría utilizarse para potencialmente identificar al paciente⁴. En ese sentido, el uso de filtros que separen el ADN humano del ADN de los microorganismos podría ser una manera eficiente de evitar esta situación en el futuro⁴.

Actualmente el cielo es el límite, con la NGS hoy hay un número de aplicaciones que se están investigando en el mundo. Por ejemplo: se puede realizar una transcriptómica global, en que el patrón de expresión de la bacteria en distintas condiciones; se puede identificar la microbiota para entender si la diversidad de ésta se asocia a enfermedad; también se puede estudiar el patrón de metilación (epigenética)⁴. Cada día vemos nuevas investigaciones y nueva evidencia de que estas tecnologías están revolucionando el modo de realizar investigaciones epidemiológicas.

Aplicabilidad y limitaciones de la genómica llevada al uso clínico en Chile

De acuerdo al mapa mundial de presencia de equipos de secuenciación masiva en los diversos países, realizado por el investigador James Hadfield de la Universidad de Cambridge (<http://omicsmaps.com/>), en Chile pueden encontrarse secuenciadores en Santiago, Concepción y Puerto Montt. Sin embargo, entre las limitantes de la real aplicabilidad de la NGS al uso clínico en Chile están:

- Acceso a la tecnología, ya que estos equipos están principalmente en Centros de Genómica o en Universidades.
- El costo, ya que a pesar que el costo de secuenciar ha disminuido de forma extraordinaria, en nuestro país la importación de los reactivos y kits de secuenciación agrega un valor considerable al precio final. Por ejemplo, el costo de la secuenciación de un genoma bacteriano con el MiSeq™ de Illumina en Chile es de aproximadamente \$250 USD versus \$100 USD en Estados Unidos (cotizaciones personales realizadas en Chile y en E.U.A.). Y finalmente
- La falta de expertos profesionales que puedan analizar e interpretar los datos generados, siendo esta limitante, un problema global.

Este es una ciencia que, al parecer, evoluciona más rápido que el entrenamiento de profesionales del área de



la salud. La aplicación de las NGS va a requerir de innovación en muchos aspectos, ya que requiere un cambio desde, por ejemplo “identificar” el perfil de resistencia antimicrobiana en una placa de agar con sensidiscos de antimicrobianos, a “predecir” el perfil de resistencia del agente, basándose en algoritmos que identifiquen la presencia de genes de resistencia (ej. Programa ResFinder 2.1, disponible en <https://cge.cbs.dtu.dk/services/ResFinder/>). Probablemente en nuestro país para que estas tecnología se usen rutinariamente para salvar vidas, se necesitará de una estrecha colaboración entre las Universidades, el Sector Privado y el Gobierno; modelo que ha sido exitoso en países desarrollados, donde médicos en hospitales, junto con investigadores en la Academia y del Gobierno, han trabajado y publicado en conjunto aplicaciones de NGS en la clínica³¹.

Conclusiones

No puede negarse que la NGS están revolucionado el campo del diagnóstico y epidemiología de enfermedades infecciosas. La disminución en los costos de secuenciación ha determinado que progresivamente se estén incorporando como herramienta de análisis epidemiológicos y de diagnóstico, especialmente en países desarrollados. A partir de esto, se están generando proyectos que permitirán almacenar gran cantidad de secuencias de agentes infecciosos, con especial énfasis en agentes zoonóticos transmitidos por los alimentos. Esta información permitirá beneficios como diagnóstico en menor tiempo, y desarrollo de técnicas que permitan determinar el origen de un brote, o diagnóstico de enfermedades nosocomiales. Especial utilidad tendrá, sobre todo, en casos que las metodologías tradicionales como microbiología clásica no son de mucha utilidad. Disciplinas como la metagenómica serán herramientas esenciales en el futuro, ya que al ser independientes del cultivo permiten detectar agentes causales de enfermedades emergentes desconocidas. Importante a considerar es que en la medida que se implementan estas disciplinas también se consideren los asuntos éticos relacionados con la información privada del paciente.

A medida que pasa el tiempo, se van actualizando los conocimientos en estas tecnologías y existe mayor disponibilidad de herramientas computacionales y analíticas para lidiar con la gran cantidad de información generada a partir de estas tecnologías. El “big data” que ya está disponible y el que en el futuro estará en la red. Ahora es importante concentrar los esfuerzos en entrenar personal en la medida que se vayan adquiriendo estas tecnologías. Es importante que sepan afrontar esta gran cantidad de información ya que parece ser la piedra angular de la epidemiología moderna.

Agradecimientos. Agradecemos el financiamiento del Proyecto Fondecyt de iniciación 11140108 a nombre de Andrea Moreno Switt. Viviana Toledo es financiada por la Vicerrectoría de Investigación y Doctorado de la Universidad Andrés Bello.

Resumen

La nueva generación de secuenciadores (NGS) ha llegado para cambiar el modo de realizar investigación. Particularmente en Infectología, estas tecnologías modernas la han dirigido a una próxima era, denominada la “era de la genómica”. En países desarrollados, las NGS se están utilizando en un gran número de aplicaciones, las que incluyen diagnóstico clínico, epidemiología y microbiología. En la actualidad, secuenciar el genoma de un microorganismo completo, ya sea bacteriano o viral, cuesta aproximadamente \$100 dólares, precio bastante asequible en comparación a los precios de las tecnologías más antiguas. En esta revisión se describen algunas publicaciones recientes que han utilizado la secuenciación de genomas completos para, (i) rastrear brotes de enfermedades transmitidas por alimentos, (ii) la preparación de base de datos de genomas para los Gobiernos (iii) investigación de enfermedades nosocomiales, y (iv) en diagnóstico clínico. La era de la genómica está aquí, y llegó para quedarse, por lo cual se debe enfocar todos los esfuerzos en aprender a utilizar la gran cantidad de “datos masivos” generados por estas tecnologías, para reducir el impacto de las enfermedades infecciosas y así, mejorar la salud de personas y animales.

Referencias bibliográficas

- 1.- Pareek C S, Smoczynski R, Tretyn A. Sequencing technologies and genome sequencing. *J Appl Genet* 2011; 52: 413-35.
- 2.- Xuan J, Yu Y, Qing T, Guo L, Shi L. Next-generation sequencing in the clinic: promises and challenges. *Cancer Lett* 2013; 340: 284-95.
- 3.- Buermans H P, den Dunnen J T. Next generation sequencing technology: Advances and applications. *Biochim Biophys Acta* 2014; 1842: 1932-45.
- 4.- Bergholz T M, Moreno Switt A I, Wiedmann M. Omics approaches in food safety: fulfilling the promise? *Trends Microbiol* 2014; 22: 275-81.
- 5.- Shendure J, Ji H. Next-generation DNA sequencing. *Nat Biotechnol* 2008; 26: 1135-45.
- 6.- Wang Y, Yang Q, Wang Z. The evolution of nanopore sequencing. *Front Genet* 2015; 5: 449.
- 7.- Check Hayden E. Data from pocket-sized genome sequencer unveiled. *Nature* 2014 Feb 14; Available from: <http://www.nature.com/news/data-from-pocket-sized-genome-sequencer-unveiled-1.14724>.
- 8.- Köser C U, Holden M T G, Ellington M J, Cartwright E J P, Brown N M, Ogilvy-Stuart A L, et al. Rapid whole-genome sequencing for investigation of a neonatal MRSA outbreak. *N Engl J Med* 2012; 366: 2267-75.



- 9.- Wiedmann M. Subtyping of bacterial foodborne pathogens. *Nutr Rev* 2002; 60: 201-8.
- 10.- Ranieri M L, Shi C, Moreno Switt A I, den Bakker H C, Wiedmann M. *Salmonella* serovar prediction: Comparison of typing methods with a new procedure based on sequence characterization. *J Clin Microbiol* 2013; 51: 1786-97.
- 11.- Moreno Switt A I, den Bakker H C, Vongkamjan K, Hoelzer K, Warnick L D, Cummings K J, et al. *Salmonella* bacteriophage diversity reflects host diversity on dairy farms. *Food Microbiol*. 2013; 36: 275-85.
- 12.- Deng X, Desai PT, den Bakker HC, Mikoleit M, Tolar B, Trees E, et al. Genomic epidemiology of *Salmonella enterica* serotype Enteritidis based on population structure of prevalent lineages. *Emerg Infect Dis* 2014; 20: 1481-9
- 13.- Allard M W, Luo Y, Strain E, Pettengill J, Timme R, Wang C, et al. On the evolutionary history, population genetics and diversity among isolates of *Salmonella* Enteritidis PFGE pattern JEGX01.0004. *PLoS One* 2013; 8: e55254.
- 14.- den Bakker H C, Allard M W, Bopp D, Brown E W, Fontana J, Iqbal Z, et al. Rapid whole-genome sequencing for surveillance of *Salmonella enterica* serovar Enteritidis. *Emerg Infect Dis* 2014; 20: 1306-14.
- 15.- Public Health England. Whole genome sequencing to revolutionise investigation of outbreaks of infectious disease. Press release 2014; Available from: <https://www.gov.uk/government/news/whole-genome-sequencing-to-revolutionise-investigation-of-outbreaks-of-infectious-disease>
- 16.- Toro M, Retamal P, Allard M, Brown E W, Evans P, González-Escalona N. Draft genome sequences of 33 *Salmonella enterica* clinical and wildlife isolates from Chile. *Genome Announc* 2015; 3: 2.
- 17.- Kundu S, Lockwood J, Depledge D P, Chaudhry Y, Aston A, Rao K, et al. Next-generation whole genome sequencing identifies the direction of norovirus transmission in linked patients. *Clin Infect Dis* 2013; 57: 407-14.
- 18.- Mathijs E, Denayer S, Palmeira L, Botteldoorn N, Scipioni A, Vanderplasschen A, et al. Novel norovirus recombinants and of GII.4 sub-lineages associated with outbreaks between 2006 and 2010 in Belgium. *Viol J* 2011;8: 310.
- 19.- Baillie G J, Galiano M, Agapow P M, Myers R, Chiam R, Gall A, et al. Evolutionary dynamics of local pandemic H1N1/2009 influenza virus lineages revealed by whole-genome analysis. *J Virol* 2012; 86: 11-8.
- 20.- Gire S K, Goba A, Andersen K G, Sealfon R S G, Park D J, Kanneh L, et al. Genomic surveillance elucidates Ebola virus origin and transmission during the 2014 outbreak. *Science* 2014; 345: 1369-72.
- 21.- Mellmann A, Harmsen D, Cummings C A, Zentz E B, Leopold S R, Rico A, et al. Prospective genomic characterization of the german enterohemorrhagic *Escherichia coli* O104:H4 outbreak by rapid next generation sequencing technology. Ahmed N, editor. *PLoS One* 2011; 6: e22751.
- 22.- Ho C-C, Yuen K-Y, Lau S K P, Woo P C Y. Rapid identification and validation of specific molecular targets for detection of *Escherichia coli* O104:H4 outbreak strain by use of high-throughput sequencing data from nine genomes. *J Clin Microbiol* 2011; 49: 3714-6.
- 23.- Pritchard L, Holden N J, Bielaszewska M, Karch H, Toth I K. Alignment-free design of highly discriminatory diagnostic primer sets for *Escherichia coli* O104:H4 outbreak strains. *PLoS One* 2012;7:e34498.
- 24.- Culligan E P, Sleator R D, Marchesi J R, Hill C. Metagenomics and novel gene discovery: Promise and potential for novel therapeutics. *Virulence* 2014; 5: 399-412.
- 25.- Canuti M, van der Hoek L. Virus discovery: are we scientists or genome collectors? *Trends Microbiol* 2014; 22: 229-31.
- 26.- Fischer N, Rohde H, Indenbirken D, Günther T, Reumann K, Lütgehetmann M, et al. Rapid metagenomic diagnostics for suspected outbreak of severe pneumonia. *Emerg Infect Dis* 2014;20: 1072-5.
- 27.- Nakamura S, Maeda N, Miron I M, Yoh M, Izutsu K, Kataoka C, et al. Metagenomic diagnosis of bacterial infections. *Emerg Infect Dis* 2008; 14: 1784-6.
- 28.- Kawai T, Sekizuka T, Yahata Y, Kuroda M, Kumeda Y, Iijima Y, et al. Identification of *Kudoa septempunctata* as the causative agent of novel food poisoning outbreaks in Japan by consumption of *Paralichthys olivaceus* in raw fish. *Clin Infect Dis* 2012; 54: 1046-52.
- 29.- Pallen M J. Diagnostic metagenomics: potential applications to bacterial, viral and parasitic infections. *Parasitology* 2014; 141: 1856-62.
- 30.- Scallan E, Griffin P M, Angulo F J, Tauxe R V, Hoekstra R M. Foodborne illness acquired in the United States-unspecified agents. *Emerg Infect Dis* 2011;17: 16-22.
- 31.- Weinstock G M, Peacock S J. Next-generation pathogen genomics. *Genome Biol* 2014; 15: 528.